MAS223 Statistical Inference and Modelling

Exercises and Solutions

The exercises are grouped into sections, corresponding to chapters of the lecture notes. Within each section exercises are divided into warm-up questions, ordinary questions, and challenge questions. Note that there are no exercises accompanying Chapter 8.

The vast majority of exercises are ordinary questions. Ordinary questions will be used in homeworks and tutorials; they cover the material content of the course. Warm-up questions are typically easier, often nothing more than revision of relevant material from first year courses. Challenge questions are typically harder and test ingenuity.

This version of the exercises also contains solutions, which are written in blue. Solutions to challenge questions are not always included, hints may be given instead. Some of the solutions mention common pitfalls, written in red, which are mistakes that are (sometimes) easily made.

The solutions sometimes omit intermediate steps of basic calculations, which are left to the reader. For example, they may simply state $\int_0^x \lambda e^{-\lambda u} = 1 - e^{-\lambda u}$, and leave you to fill in the intermediate steps.

Contents

1 Univariate Distribution Theory

Warm-up Questions

1.1 Let X be a random variable taking values in $\{1, 2, 3\}$, with $\mathbb{P}[X = 1] = \mathbb{P}[X = 2] = 0.4$. Find $\mathbb{P}[X = 3]$, and calculate both $\mathbb{E}[X]$ and $\text{Var}[X]$.

Solution. Since $\mathbb{P}[X = 1] + \mathbb{P}[X = 2] + \mathbb{P}[X = 3] = 1$, we have $\mathbb{P}[X = 3] = 0.2$. With this, we can calculate

$$\mathbb{E}[X] = 1\mathbb{P}[X = 1] + 2\mathbb{P}[X = 2] + 3\mathbb{P}[X + 3] = 1.8$$
$$\mathbb{E}[X^2] = 1^2\mathbb{P}[X = 1] + 2^2\mathbb{P}[X = 2] + 3^2\mathbb{P}[X + 3] = 3.8$$

Using that $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, we have $\text{Var}(X) = 0.56$.

1.2 Let Y be a random variable with probability density function (p.d.f.) $f(y)$ given by

$$f(y) = \begin{cases} y/2 & \text{for } 0 \le y < 2; \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability that Y is between $\frac{1}{2}$ and 1. Calculate $\mathbb{E}[Y]$ and $\text{Var}[Y]$.

Solution. We have $\mathbb{P}[Y \in [\frac{1}{2}, 1]] = \int_{1/2}^{1} (y/2)\, dy = 3/16$. Similarly,

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} y f(y) = \int_0^2 y(y/2)\, dy = 4/3$$
$$\mathbb{E}[Y^2] = \int_0^2 (y^3/2)\, dy = 2$$

so $\text{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = 2/9$.

Ordinary Questions

1.3 Define $F : \mathbb{R} \to [0, 1]$ by

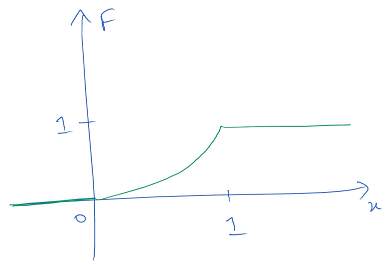
$$F(y) = \begin{cases} 0 & \text{for } y \le 0; \\ y^2 & \text{for } y \in (0, 1); \\ 1 & \text{for } y \ge 1. \end{cases}$$

(a) Sketch the function F, and check that it is a distribution function.

(b) If Y is a random variable with distribution function F, calculate the p.d.f. of Y.

Solution.

(a) Sketch of F should look like



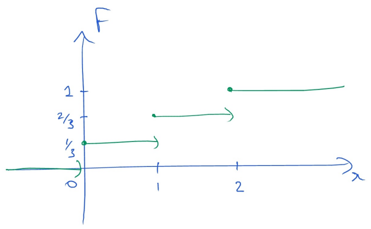
From the graph, F is continuous and (non-strictly) increasing. We have $F(x) = 0$ for all $x \le 0$, so $\lim_{x\to-\infty} F(x) = 0$. Similarly, $F(x) = 1$ for all $x \ge 1$, so $\lim_{x\to\infty} F(x) = 1$. Hence, F satisfies all the properties of a distribution function.

(b) We have $f(y) = F'(y)$, so treating each case in turn,

$$f(y) = \begin{cases} 0 & \text{for } y \le 0; \\ 2y & \text{for } y \in (0,1); \\ 0 & \text{for } y \ge 1. \end{cases}$$

1.4 Let X be a discrete random variable, taking values in $\{0, 1, 2\}$, where $\mathbb{P}[X = n] = \frac{1}{3}$ for $n \in \{0, 1, 2\}$. Sketch the distribution function $F_X : \mathbb{R} \to \mathbb{R}$.

Solution. Sketch should look like



Pitfall: The graph of F is not continuous; it jumps at $0, 1, 2$, and is otherwise constant.

1.5 Define $f : \mathbb{R} \to [0, 1]$ by

$$f(x) = \begin{cases} 0 & \text{for } x < 0; \\ e^{-x} & \text{for } x \ge 0. \end{cases}$$

(a) Show that f is a probability density function.

(b) Find the corresponding distribution function and evaluate $\mathbb{P}[1 < X < 2]$.

Solution.

(a) Clearly $f(x) \ge 0$ for all x, and

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{\infty} e^{-x}\,dx = \left[-e^{-x}\right]_0^{\infty} = 1,$$

so f is a probability density function.

(b) We need to calculate $F(x) = \mathbb{P}[X \le x] = \int_{-\infty}^{x} f(u)\,du$. For $x \le 0$ we have $F(x) = \int_{-\infty}^{x} 0\,dx = 0$. For $x \ge 0$, we have $F(x) = \int_{-\infty}^{0} 0\,du + \int_0^x e^{-u}, du = 0 + (1 - e^{-x})$. Thus,

$$F(x) = \begin{cases} 0 & \text{for } x \le 0; \\ 1 - e^{-x} & \text{for } x \ge 0. \end{cases}$$

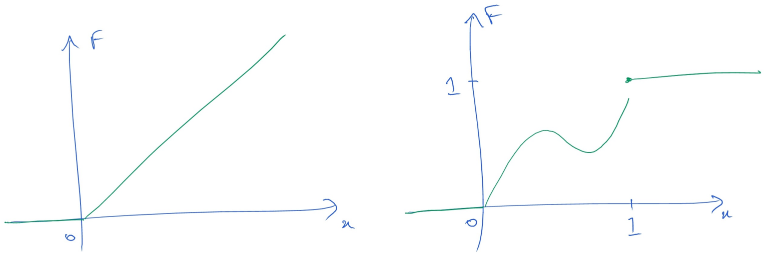
Hence, $\mathbb{P}[1 < X < 2] = \mathbb{P}[X < 2] - \mathbb{P}[X \le 1] = \mathbb{P}[X \le 2] - \mathbb{P}[X \le 1] = F(2) - F(1) = e^{-1} - e^{-2}$.

1.6 Sketch graphs of each of the following two functions, and explain why each of them is not a distribution function.

(a) $F(x) = \begin{cases} 0 & \text{for } x \le 0; \\ x & \text{for } x > 0. \end{cases}$

(b) $F(x) = \begin{cases} 0 & \text{for } x < 0; \\ x + \frac{1}{4}\sin 2\pi x & \text{for } 0 \le x < 1; \\ 1 & \text{for } x \ge 1. \end{cases}$

Solution. Sketches should look like



For (a), $F(x) > 1$ for $x > 1$, so F does not stay between 0 and 1. For (b), for $x \in [0, 1)$ we have $F'(x) = f(x) = 1 + \frac{2\pi}{4}\cos(2\pi x)$, which is negative at, for example, $x = \frac{1}{2}$, so (as is clear from the graph) F is not an increasing function.

1.7 Let $k \in \mathbb{R}$ and define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} k(x - x^2) & \text{for } x \in (0, 1), \\ 0 & \text{otherwise.} \end{cases}$$

Find the value of k for which $f(x)$ is a probability density function, and calculate the probability that X is greater than $\frac{1}{2}$.

Solution. We need $f(x) \ge 0$ for all x, so we need $k \ge 0$. Also, we need

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = k\int_0^1 x - x^2\,dx = k\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{k}{6}.$$

So, $k = 6$. Therefore,

$$\mathbb{P}[X \ge \tfrac{1}{2}] = \int_{\frac{1}{2}}^{\infty} f(x)\,dx = \int_{\frac{1}{2}}^{1} 6(x - x^2)\,dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_{\frac{1}{2}}^{1} = \frac{1}{2}.$$

1.8 The probability density function $f(x)$ is given by

$$f(x) = \begin{cases} 1+x & \text{for } -1 \le x \le 0; \\ 1-x & \text{for } 0 < x < 1; \\ 0 & \text{otherwise.} \end{cases}$$

Find the corresponding distribution function $F(x)$ for all real x.

Solution. We have $F(-1) = \mathbb{P}[X \le -1] = P[X < -1] = 0$. So $F(x) = 0$ for all $x \le -1$. If $x \in [-1, 0]$ then

$$F(x) = \int_{-\infty}^{x} f(u)\,du = F(-1) + \int_{-1}^{x} (1+u)\,du = 0 + \frac{(x+1)^2}{2}.$$

Now, for $x \in [0, 1]$, we have

$$F(x) = \int_{-\infty}^{x} f(u)\,du = F(0) + \int_{0}^{x} (1-u)\,du = \frac{1}{2} + x - \frac{x^2}{2} = \frac{1+2x-x^2}{2}.$$

Pitfall: Forgetting the F(0) results in missing out the term $\frac{1}{2}$. It needs to be present because for $x \in (0, 1)$ we have

$$F(x) = \mathbb{P}[X \le x] = \mathbb{P}[X \le 0] + \mathbb{P}[0 < X < x] = F(0) + \int_{0}^{x} (1-u)\,du.$$

Note that in the case of $x \in [-1, 0]$ the equivalent term was $F(-1)$ and was equal to 0.

Therefore, we have $F(1) = 1$. Since F is increasing and must stay between 0 and 1, we have $F(x) = 1$ for all $x \ge 1$.

Thus the distribution function $F(x)$ is

$$F(x) = \begin{cases} 0, & \text{for } x < -1 \\ \frac{(x+1)^2}{2}, & \text{for } -1 \le x < 0 \\ \frac{1+2x-x^2}{2}, & \text{for } 0 \le x < 1 \\ 1, & \text{for } x \ge 1 \end{cases}$$

1.9 Let

$$F(x) = \frac{e^x}{1+e^x} \quad \text{for all real } x.$$

(a) Show that F is a distribution function, and find the corresponding p.d.f. f.

(b) Show that $f(-x) = f(x)$.

(c) If X is a random variable with this distribution, evaluate $\mathbb{P}[|X| > 2]$.

Solution.

(a) Since $e^x \to 0$ as $x \to \infty$, we have

$$\lim_{x\to-\infty} \frac{e^x}{1+e^x} = \frac{0}{1+0} = 0$$
$$\lim_{x\to\infty} \frac{e^x}{1+e^x} = \lim_{x\to\infty} \frac{1}{e^{-x}+1} = \frac{1}{0+1} = 1$$

Pitfall: It does not make sense to use that $\lim_{x\to\infty} e^x = \infty$ and then incorrectly claim that $\frac{\infty}{1+\infty} = 1$.

Using the quotient rule, the derivative of f, the corresponding p.d.f. , is

$$f(x) = \frac{e^x(1+e^x) - e^x e^x}{(1+e^x)^2} = \frac{e^x}{(1+e^x)^2}.$$

Therefore, $f(x) > 0$ for all $x \in \mathbb{R}$, so F is an increasing function.

Since $x \mapsto e^x$ is continuous, F is a composition of sums and (non-zero) divisions of continuous functions; therefore F is continuous.

Hence, F satisfies all the properties of a distribution function.

(b) $f(-x) = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{e^{2x}e^{-x}}{e^{2x}(1+e^{-x})^2} = \frac{e^x}{(e^x+1)^2} = f(x)$.

(c) We have

$$\begin{aligned}\mathbb{P}[|X| > 2] &= \mathbb{P}[X < -2] + \mathbb{P}[X > 2] \\ &= \mathbb{P}[X \le -2] + (1 - \mathbb{P}[X \le 2]) \\ &= 1 + F(-2) - F(2) \\ &= 1 + \tfrac{e^{-2}}{1+e^{-2}} - \tfrac{e^2}{1+e^2} \approx 0.238.\end{aligned}$$

Note that here we used $\mathbb{P}[X < -2] = \mathbb{P}[X \le -2]$, which holds because F is continuous.

1.10 Show that $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$, defined for all $x \in \mathbb{R}$, is a probability density function.

Solution. Clearly $f(x) \ge 0$ for all x, and

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{1+x^2}\,dx = \frac{1}{\pi}\Big[\arctan(x)\Big]_{-\infty}^{\infty} = \frac{1}{\pi}\left(\frac{\pi}{2} - \frac{-\pi}{2}\right) = 1.$$

1.11 (a) For which values of $r \in [0, \infty)$ is $\int_1^\infty x^{-r}\,dx$ finite?

(b) Show that

$$f(x) = \begin{cases} x^{-2} & \text{if } x > 1 \\ 0 & \text{otherwise} \end{cases}$$

is a probability density function.

(c) Show that the expectation of a random variable X, with the probability density function f given in (b), is not defined.

(d) Give an example of a random variable Y for which $\mathbb{E}[Y] < \infty$ but $\mathbb{E}[Y^2]$ is not defined.

Solution.

(a) For $r \neq -1$ we have

$$\int_1^\infty x^{-r}dx = \lim_{n\to\infty}\int_1^n x^{-r}dx = \lim_{n\to\infty}\frac{n^{-r+1}}{-r+1} - \frac{1}{-r+1}$$

which is finite (and equal to $\frac{1}{r-1}$) if $-r+1 < 0$ and infinite if $-r+1 > 0$. When $r = -1$ we have

$$\int_1^\infty x^{-1}dx = \lim_{n\to\infty}\int_1^n x^{-1}dx = \lim_{n\to\infty}\log n = \infty$$

Hence, $\int_1^\infty x^{-r}dx$ is finite if and only if $r > 1$.

(b) Clearly $f(x) \geq 0$ for all x and

$$\int_1^\infty x^{-2} dx = \left[(-x^{-1})\right]_{x=1}^\infty = 1$$

so f is a probability density function.

(c) we have $\int_{-\infty}^\infty x f_X(x) = \int_1^\infty x^{-1}\, dx$, which by part (a) is infinite. Hence the expectation of X is not defined.

(d) Let

$$f(y) = \begin{cases} 2y^{-3} & \text{if } x > 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then $f(y) \geq 0$ for all y, and

$$\int_{-\infty}^\infty f(y)\, dy = \int_1^\infty 2y^{-3}\, dy = \left[-y^{-2}\right]_{y=1}^\infty = 1$$

so f is a probability density function. From (a) we have that

$$\int_{-\infty}^\infty y^2 f_Y(y) dy = 2\int_1^\infty y^{-1} dy$$

is infinite and that

$$\int_{-\infty}^\infty y f_Y(y) dy = 2\int_1^\infty y^{-2} dy = \frac{2}{3}$$

is finite. Hence, if Y is a random variable with p.d.f. f, then $\mathbb{E}[Y]$ is defined but $\mathbb{E}[Y^2]$ is not.

1.12 The discrete random variable X has the probability function

$$\mathbb{P}[X = x] = \begin{cases} \frac{1}{x(x+1)} & \text{for } x \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

(a) Use the partial fractions of $\frac{1}{x(x+1)}$ to show that $\mathbb{P}[X \leq x] = 1 - \frac{1}{x+1}$, for all $x \in \mathbb{N}$.

(b) Write down the distribution function $F(x)$ of X, for $x \in \mathbb{R}$. Sketch its graph. What are the values of $F(2)$ and $F(\frac{3}{2})$?

(c) Evaluate $\mathbb{P}[10 \leq X \leq 20]$.

(d) Is $\mathbb{E}[X]$ defined? If so, what is $\mathbb{E}[X]$? If not, why not?

Solution.

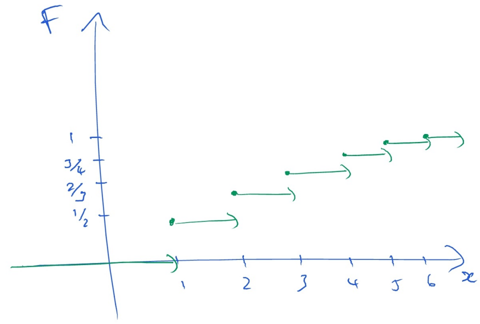
(a) Using partial fractions we obtain the identity $\frac{1}{x(x+1)} = \frac{1}{x} - \frac{1}{x+1}$, provided $x \neq 0, -1$. Hence, if $x \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}[X \leq x] &= \sum_{i=1}^{x} \mathbb{P}[X = x] \\ &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \ldots + \left(\frac{1}{x} - \frac{1}{x+1}\right) \\ &= 1 - \frac{1}{x+1}.\end{aligned}$$

(b) The distribution function is

$$F(u) = \begin{cases} 1 - \frac{1}{x+1}, & \text{for } u \in [x, x+1) \text{ where } x \in \mathbb{N} \\ 0, & \text{for } u < 1 \end{cases}$$

which looks like



$F(2) = 1 - \frac{1}{3} = \frac{2}{3}$ and $F(\frac{3}{2}) = F(1) = \frac{1}{2}$.

Pitfall: The graph of F is not continuous. The formula obtained in part (a) is only valid for $x \in \mathbb{N}$, and not for all $x \in \mathbb{R}$. Since X is a discrete random variable, its distribution function jumps at the points where X takes values (i.e. at $x \in \mathbb{N}$) and is constant in between those points.

(c) Since X is discrete,

$$\mathbb{P}[10 \leq X \leq 20] = \mathbb{P}[X \leq 20] - \mathbb{P}[X \leq 9] = F(20) - F(9) = \tfrac{11}{210}.$$

Pitfall: X is discrete, and $\mathbb{P}[X = 10] > 0$. So it's $F(20) - F(9)$, and not $F(20) - F(10)$.

(d) Since X is discrete, $\mathbb{E}[X]$ is defined if and only if $\sum_x x\mathbb{P}[X = x]$ converges. In our case, this sum is equal to

$$\sum_{x=1}^{\infty} x \frac{1}{x(x+1)} = \sum_{x=1}^{\infty} \frac{1}{x+1}$$

which diverges.

To see that the sum diverges, we can write it as

$$\left(\frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots$$

and note that each bracketed term is at least $\frac{1}{2}$; of course $\sum_{x=1}^{\infty} \frac{1}{2} = \infty$.

Challenge Questions

1.13 Show that there is no random variable X, with range $\mathbb{N}$, such that $\mathbb{P}[X = n]$ is constant for all $n \in \mathbb{N}$.

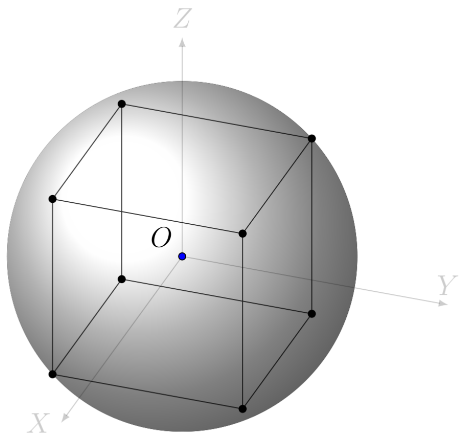
Solution. Since X has range $\mathbb{N}$ we have $\mathbb{P}[X \in \mathbb{N}] = 1$. If $\mathbb{P}[X = n] = \mathbb{P}[X = 1]$ for all n then we would have

$$1 = \mathbb{P}[X \in \mathbb{N}] = \sum_{n=1}^{\infty} \mathbb{P}[X = n] = \sum_{n=1}^{\infty} \mathbb{P}[X = 1].$$

If $\mathbb{P}[X = 1] = 0$ then $\mathbb{P}[X \in \mathbb{N}] = 0$, which is impossible, but similarly if $\mathbb{P}[X = 1] > 0$ then the sum is equal to $+\infty$, which is not possible either.

Hence, no such random variable exists.

1.14 Recall the meaning of 'inscribing' a cube within a sphere: the cube sits inside of the sphere, with each vertex of the cube positioned on the surface of the sphere. It is not especially easy to illustrate this on a two dimensional page, but here is an attempt:



Suppose that ten percent of the surface of the sphere is coloured blue, and the rest of the surface is coloured red. Show that, regardless of which parts are coloured blue, it is always possible to inscribe a cube within the sphere in such a way as all vertices of the cube are red.

Hint: A cube has eight corners. Suppose that position of the cube is sampled uniformly from the set of possible positions. What is the expected number of corners that are red?

Solution. Let X be a cube inscribed within the sphere, orientated uniformly at random. Strictly speaking, in three-dimensional spherical coordinates (r, θ, ϕ) this means we let the polar angle θ and azimuth angle ϕ be independent uniform random variables on $(0, 2\pi)$. More importantly, it means that the location of a given vertex of the cube is distributed uniformly on the surface of the sphere.

Let X be the number of corners that are red. Label the corners from $i = 1, \ldots, 8$. We can write

$$X = \sum_{i=1}^{8} \mathbb{1}_{\{A_i = \text{red}\}},$$

where A_i is the colour of the i^{th} corner. Here $\mathbb{1}_{\{A_i=\text{red}\}}$ is equal to 1 if A_i is red and equal to zero if A_i is blue. Hence,

$$\mathbb{E}[X] = \sum_{i=1}^{8} \mathbb{E}\left[\mathbb{1}_{\{A_i=\text{red}\}}\right] = \sum_{i=1}^{8} \mathbb{P}[A_i = \text{red}].$$

Since A_i is uniformly distributed on the surface of the sphere, and 90% of the sphere is red, we have $\mathbb{P}[A_i = \text{red}] = \frac{9}{10}$. Hence,

$$\mathbb{E}[X] = \tfrac{9}{10} \times 8 = 7.2.$$

Note that X can only take the values $\{1, 2, \ldots, 8\}$. Since $\mathbb{E}[X] > 7$, we must have $\mathbb{P}[X = 8] > 0$. Therefore, there are orientations of the cube for which all 8 vertices are red.

2 Standard Univariate Distributions

Warm-up Questions

2.1 (a) A standard fair dice is rolled 5 times. Let X be the number of sixes rolled. Which distribution (and which parameters) would you use to model X?

(b) A fair coin is flipped until the first head is shown. Let X be the total number of flips, including the final flip on which the first head appears. Which distribution (and which parameter) would you use to model X?

Solution.

(a) The binomial distribution, with parameters $n = 5$ and $p = \frac{1}{6}$.

(b) The geometric distribution, with parameter $p = \frac{1}{2}$.

Ordinary Questions

2.2 Let $\lambda > 0$. Write down the p.d.f. f of the random variable X, where $X \sim Exp(\lambda)$, and calculate its distribution function F. Hence, show that $\frac{f(t)}{1-F(t)}$ is constant for $t > 0$.

Solution. The p.d.f. of X is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x > 0; \\ 0 & \text{otherwise.} \end{cases}$$

It's distribution function $F(x) = \int_{-\infty}^{x} f(u)\,du$ is clearly zero for $x \leq 0$, and for $x > 0$ we have $\int_{-\infty}^{x} f(u)\,du = \int_{0}^{x} \lambda e^{-\lambda u}\,du = 1 - e^{-\lambda x}$. Therefore,

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Hence, for $t > 0$ we have $\frac{f(t)}{1-F(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$.

Pitfall: Don't forget the 'otherwise' case where $f(x) = 0$ or $F(x) = 0$. The same comment applies to many other questions.

2.3 Let $\lambda > 0$ and let X be a random variable with $Exp(\lambda)$ distribution. Let $Z = \lfloor X \rfloor$, that is let Z be X rounded down to the nearest integer. Show that Z is geometrically distributed with parameter $p = 1 - e^{-\lambda}$.

Solution. Since $X > 0$, we have $Z \in \{0, 1, 2, \ldots\}$, hence $\mathbb{P}[Z = z] = 0$ for all other z. For $n \in \{0, 1, 2, \ldots\}$ we have

$$\begin{aligned} \mathbb{P}[Z = n] &= \mathbb{P}[n \leq X < n + 1] \\ &= \int_{n}^{n+1} \lambda e^{-\lambda x}\,dx \\ &= e^{-\lambda n} - e^{-\lambda(n+1)} \\ &= e^{-\lambda n}(1 - e^{-\lambda}) \\ &= (1 - p)^n p. \end{aligned}$$

which is the probability function of the geometric distribution.

2.4 Let $\mu \in \mathbb{R}$. Let X_1 and X_2 be independent random variables with distributions $N(\mu, 1)$ and $N(\mu, 4)$, respectively. Let T_1, T_2 and T_3 be defined by

$$T_1 = \frac{X_1 + X_2}{2}, \qquad T_2 = 2X_1 - X_2, \qquad T_3 = \frac{4X_1 + X_2}{5}.$$

Find the mean and variance of T_1, T_2 and T_3. Which of $\mathbb{E}[T_1]$, $\mathbb{E}[T_2]$ and $\mathbb{E}[T_3]$ would you prefer to use as an estimator of μ?

Solution. We have $\mathbb{E}[T_1] = \frac{1}{2}(\mathbb{E}[X_1] + \mathbb{E}[X_2]) = \mathbb{E}[T_2] = \mu$. Similarly, $\mathbb{E}[T_2] = \mathbb{E}[T_3] = \mu$, so all are unbiased when used as estimators of μ. We have

$$\mathrm{Var}(T_1) = \left(\tfrac{1}{2}\right)^2 (\mathrm{Var}(X_1) + 2\,\mathrm{Cov}(X_1, X_2) + \mathrm{Var}(X_2)) = \tfrac{1}{4}(1 + 0 + 4) = \tfrac{5}{4},$$

and similarly $\mathrm{Var}(T_2) = 8$, $\mathrm{Var}(T_3) = \frac{4}{5}$.

On this information, we prefer $\mathbb{E}[T_3]$ as an estimator of μ, because T_3 has the smallest variance and so is likely to be closest to its mean.

2.5 Let X be a random variable with $Ga(\alpha, \beta)$ distribution.

(a) Let $k \in \mathbb{N}$. Show that

$$\mathbb{E}[X^k] = \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k}.$$

Hence, calculate $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathrm{Var}(X)$ and verify that these formulas match the ones given in lectures.

(b) Show that $\mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{2}{\sqrt{\alpha}}$.

Solution.

(a) From the p.d.f. of the Gamma distribution, we have

$$\begin{aligned}
\mathbb{E}[X^k] &= \int_0^\infty x^k \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\, dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{k+\alpha-1} e^{-\beta x}\, dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(k+\alpha)}{\beta^{k+\alpha}} \\
&= \frac{\Gamma(k+\alpha)}{\beta^k \Gamma(\alpha)} \\
&= \frac{(k+\alpha-1)(k+\alpha-2)\cdots(k+\alpha-k+1)(k+\alpha-k)\Gamma(\alpha)}{\beta^k \Gamma(\alpha)} \\
&= \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k}.
\end{aligned}$$

To deduce the second line from the third line, we use Lemma 2.3 from lecture notes, and to deduce the fifth line from the fourth line we use Lemma 2.2, also from lecture notes.

Pitfall: It is true that $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$, but if $n \notin \mathbb{N}$ then $(n-1)!$ does not make sense. For general $\alpha \in (1, \infty)$ we have $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$, which is what must be be used to deduce the fifth line above.

If $k = 1$ then $\mathbb{E}[X] = \frac{a}{b}$. If $k = 2$ then $\mathbb{E}[X^2] = \frac{\alpha(\alpha+1)}{\beta^2}$ and so the variance of X is $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{\alpha(\alpha+1)}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}$.

(b) With $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathrm{Var}(X)$, multiplying out and using the formulae from (a), we have

$$\begin{aligned}
\frac{\mathbb{E}[(X-\mu)^3]}{\sigma^3} &= \frac{\mathbb{E}[X^3 - 3X^2\mu + 3X\mu^2 - \mu^3]}{\sigma^3} \\
&= \frac{\mathbb{E}[X^3] - 3\mathbb{E}[X]\mathbb{E}[X^2] + 2\mathbb{E}[X]^3}{\sigma^3} \\
&= \frac{\frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3} - \frac{3\alpha^2(\alpha+1)}{\beta^3} + \frac{2\alpha^3}{\beta^3}}{\sigma^3} \\
&= \frac{\alpha(\alpha^2 + 2\alpha + \alpha + 2 - 3\alpha^2 - 3\alpha + 2\alpha^2)/\beta^3}{\sqrt{\alpha^3}/\beta^3} \\
&= \frac{2}{\sqrt{\alpha}}.
\end{aligned}$$

2.6 (a) Using R, you can obtain a plot of, for example, the p.d.f. of a $Ga(3, 2)$ random variable between 0 and 10 with the command

```
curve(dgamma(x,shape=3,scale=2),from=0,to=10)
```

Use R to investigate how the shape of the p.d.f. of a Gamma distribution varies with the different parameter values. In particular, fix a value of β, see how the shape changes as you vary α.

(b) Investigate the effect that changing parameters values has on the shape of the p.d.f. of the Beta distribution. To produce, for example, a plot of the p.d.f. of $Be(4, 5)$, use

```
curve(dbeta(x,shape1=4,shape2=5),from=-1,to=2)
```

Solution.

(a) You should discover that decreasing α makes the p.d.f. appear more skewed (to the right). This makes it more likely that a sample of the random variable has a large value.

(b) You should discover that the parameter `shape1` (which we normally denote by α) controls the behaviour near $x = 0$, and `shape2` (that is, β), controls the behaviour near $x = 1$. In both case, the parameters can be tuned to cause (slow or fast) explosion to ∞, convergence to 1, and (slow or fast) convergence towards 0.

2.7 Suggest which standard discrete distributions (or combination of them) we should use to model the following situations.

(a) Organisms, independently, possess a given characteristic with probability p. A sample of k organisms with the characteristic is required. How many organisms will need to be tested to achieve this sample?

(b) In Texas Hold'em Poker, players make the best hand they can by combining two cards in their hand with five 'community' cards that are placed face up on the table. At the start of the game, a player can only see their own hand. The community cards are then turned over, one by one.

A player has two hearts in her hand. Three of the community cards have been turned over, and only one of them is a heart. How many hearts will appear in the remaining two community cards?

Use a computer to find the probability of seeing $k = 0, 1, 2$ hearts.

Solution.

(a) We'll need to sample, with success probability p, until we achieve k successes. So we will need to test $N \sim NegBin(k, p)$ organisms to find our sample.

(b) There are a total of 52 cards, 13 of each of the four suits. Our player can see 5 cards, 3 of which are hearts. Therefore, the unknown cards consist of 47 cards, 10 of which are hearts. The number of hearts that will be drawn in the next two community cards is, therefore, a hypergeometric distribution with parameters $N = 47$ (population size), $k = 10$ (successes), $n = 2$ (trials). As a result (use e.g. `R`), $\mathbb{P}[X = 0] \approx 0.65$, $\mathbb{P}[X = 1] \approx 0.32$ and $\mathbb{P}[X = 2] = 0.03$.

2.8 Let X be a $N(0, 1)$ random variable. Use integration by parts to show that $\mathbb{E}[X^{n+2}] = (n + 1)\mathbb{E}[X^n]$ for any $n = 0, 1, 2, \ldots$. Hence, show that

$$\mathbb{E}[X^n] = \begin{cases} 0 & \text{if } n \text{ is odd;} \\ (1)(3)(5)\ldots(n-1) & \text{if } n \text{ is even.} \end{cases}$$

Solution. Integrating by parts, for any $n \geq 0$, gives

$$\begin{aligned} \mathbb{E}[X^n] &= \int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx \\ &= \left[\frac{1}{\sqrt{2\pi}} \frac{x^{n+1}}{n+1} e^{-x^2/2}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{x^{n+1}}{n+1} \frac{1}{\sqrt{2\pi}} (-x) e^{-x^2/2}\,dx \\ &= 0 + \int_{-\infty}^{\infty} \frac{x^{n+2}}{n+1} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx = \frac{1}{n+1}\mathbb{E}[X^{n+2}] \end{aligned}$$

Rearranging slightly,

$$\mathbb{E}[X^{n+2}] = (n + 1)\mathbb{E}[X^n].$$

Since $\mathbb{E}[X] = 0$, induction gives that $\mathbb{E}[X^n] = 0$ for all odd n. Since $\mathbb{E}[X^0] = 1$, induction gives that $\mathbb{E}[X^n] = (1)(3)(5)\ldots(n - 1)$ for all even n.

2.9 Let $X \sim N(\mu, \sigma^2)$. Show that $\mathbb{E}[e^X] = e^{\mu + \frac{\sigma^2}{2}}$.

Solution. Using the scaling properties of normal random variables from (2.2), we write $X = \mu + Y$ where $Y \sim N(0, \sigma^2)$. Hence, $e^X = e^{\mu+Y} = e^\mu e^Y$ and

$$\begin{aligned}
\mathbb{E}[e^Y] &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^y e^{-\frac{y^2}{2\sigma^2}}\,dy \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\left(y^2+2\sigma^2 y\right)\right\} dy \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\left((y+\sigma^2)^2-\sigma^4\right)\right\} dy \\
&= e^{\frac{\sigma^2}{2}}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\left\{-\frac{(y+\sigma^2)^2}{2\sigma^2}\right\} dy \\
&= e^{\frac{\sigma^2}{2}}
\end{aligned}$$

Here, we use the same method as in Example 5: in the third line we complete the square and to deduce the final line we use that the p.d.f. of a $N(-\sigma^2, \sigma)$ random variable integrates to 1. Therefore,

$$\mathbb{E}[e^X] = e^\mu \mathbb{E}[e^Y] = e^{\mu+\frac{\sigma^2}{2}}.$$

Challenge Questions

2.10 Let X be a random variable with a continuous distribution, and a strictly increasing distribution function F. Show that $F(X)$ has a uniform distribution on $(0, 1)$.

Suggest how we might use this result to simulate samples from standard distributions.

Solution. Since F is strictly increasing, it has an inverse function F^{-1}. For $x \in (0, 1)$, we have

$$\mathbb{P}[F(X) \le x] = \mathbb{P}[X \le F^{-1}(x)] = F(F^{-1}(x)) = x.$$

Hence, $F(X)$ has the uniform distribution on $(0, 1)$.

We write this as $U = F(X)$, where U is uniform on $(0, 1)$. Therefore, $F^{-1}(U) = X$. Consequently, if we can simulate uniform random variables, and calculate $F^{-1}(x)$ for given x, we can simulate X as $F^{-1}(U)$.

In fact, this is a very common way of simulating random variables. Recall that a distribution function F is not necessarily strictly increasing, but it is necessarily *non-strictly* increasing. With some care, it is possible to extend this result to cover the general case. For many standard distributions, F^{-1} can be computed explicitly.

2.11 Prove that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Hint. Thanks to the normal distribution, you know that $\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}$.

3 Transformations of Univariate Random Variables

Warm-up Questions

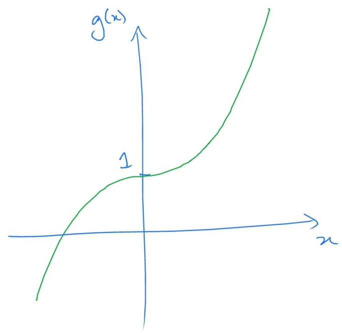
3.1 Let $g : \mathbb{R} \to \mathbb{R}$ be given by $g(x) = x^3 + 1$.

(a) Sketch the graph of g, show that g is strictly increasing, and find its inverse function g^{-1}.

(b) Let $R = [0, 2]$. Find $g(R)$.

Solution.

(a) A sketch of the function g looks like



It is clear from the graph that g is strictly increasing. If we set $y = x^3 + 1$ then $x = (y - 1)^{1/3}$, so the inverse function is $g^{-1}(y) = (y - 1)^{1/3}$.

(b) We have $g(0) = 1$ and $g(2) = 2^3 + 1 = 9$, so $g(R) = [1, 9]$.

Ordinary Questions

3.2 Let X be a random variable with p.d.f.

$$f_X(x) = \begin{cases} x^{-2} & \text{for } x > 1 \\ 0 & \text{otherwise.} \end{cases}$$

Define $g(x) = e^x$ and let $Y = g(X)$.

(a) Show that $g(x)$ is strictly increasing. Find its inverse function $g^{-1}(y)$, and $\frac{dg^{-1}(y)}{dy}$.

(b) Identify the set R_X on which $f_X(x) > 0$. Sketch g and show that $g(R_X) = (e, \infty)$.

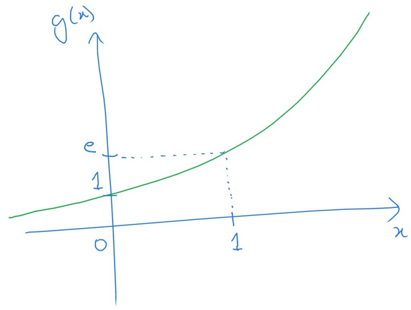
(c) Deduce from (a) and (b) that Y has p.d.f.

$$f_Y(y) = \begin{cases} (\log y)^{-2} \frac{1}{y} & \text{for } y > e \\ 0 & \text{otherwise.} \end{cases}$$

Solution.

(a) We have $\frac{dg(x)}{dx} = e^x > 0$, so g is strictly increasing. If we have $y = e^x$ then $x = \log y$, so the inverse function is $g^{-1}(y) = \log y$. Hence $\frac{dg^{-1}(y)}{dy} = \frac{1}{y}$.

(b) $f_X(x)$ is non-zero for $x \in R_X = (1, \infty)$. A sketch of g looks like



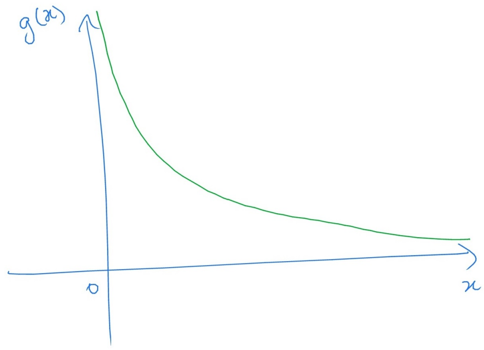
and hence $g(R_X) = (e, \infty)$.

(c) Since g is strictly increasing, we can use the formula

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \times \left|\frac{dg^{-1}(y)}{dy}\right| & \text{if } y \in g(R_X) \\ 0 & \text{otherwise,} \end{cases}$$
$$= \begin{cases} \frac{(\log y)^{-2}}{y} & \text{for } y > e \\ 0 & \text{otherwise.} \end{cases}$$

3.3 Let X be a random variable with the uniform distribution on $(0, 1)$, and let $Y = \frac{-\log X}{\lambda}$ where $\lambda > 0$. Show that Y has an $Exp(\lambda)$ distribution.

Solution. We have $f_X(x) = 1$ for $x \in (0, 1)$ (and $f_X(x) = 0$ otherwise). Hence $R_X = (0, 1)$. Our transformation is $g(x) = \frac{-\log x}{\lambda}$. For $x > 0$ we have $\frac{dg}{dx} = \frac{-1}{\lambda x} < 0$, so g is strictly decreasing. A sketch of g looks like



from which we can see that $g(R_X) = (0, \infty)$.

Writing $y = \frac{-\log x}{\lambda}$, we have $x = e^{-\lambda y}$ so $g^{-1}(y) = e^{-\lambda y}$ and $\frac{dg^{-1}}{dy} = -\lambda e^{-\lambda y}$. Hence, we can apply Lemma 3.1 to find $f_Y(y)$. Thus,

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left|\frac{dg^{-1}(y)}{dy}\right| & \text{for } y \in g(R_X) \\ 0 & \text{otherwise.} \end{cases}$$

Hence, using parts (a),(c) and (d) we obtain

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y} & \text{for } y \in (0,\infty); \\ 0 & \text{otherwise.} \end{cases}$$

This is the p.d.f. of the $Exp(\lambda)$ distribution. Hence, $Y \sim Exp(\lambda)$.

Pitfall: Don't forget to comment that g is strictly decreasing (or increasing), or else it isn't clear that you've checked whether or not Lemma 3.1 applies.

3.4 Let $\alpha, \beta > 0$.

(a) Show that $B(\alpha,\beta) = B(\beta,\alpha)$.

(b) Let X be a random variable with the $Be(\alpha,\beta)$ distribution. Show that $Y = 1 - X$ has the $Be(\beta,\alpha)$ distribution.

Solution.

(a) We have $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \frac{\Gamma(\beta)\Gamma(\alpha)}{\Gamma(\alpha+\beta)} = B(\beta,\alpha)$.

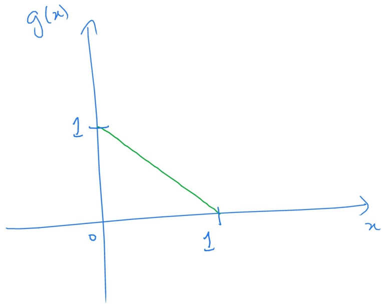
(b) We aim to use Lemma 3.1. We have

$$f_X(x) = \begin{cases} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } x \in (0,1); \\ 0 & \text{otherwise.} \end{cases}$$

Define $g : (0,1) \to (0,1)$ by

$$g(x) = 1 - x$$

and then $Y = g(X)$. Note that g is strictly increasing on $(0,1)$. We have $g^{-1}(y) = 1 - y$ and $\frac{dg^{-1}}{dy} = -1$. A sketch of g looks like



from which we can see that $g((0,1)) = (0,1)$. Hence, from Lemma 3.1 we have

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}}{dy} \right| = \frac{1}{B(\alpha,\beta)} (1-y)^{\alpha-1} y^{\beta-1} \times |-1| \,.$$

for $y \in (0,1)$, giving

$$f_Y(y) = \begin{cases} \frac{1}{B(\beta,\alpha)} y^{\beta-1}(1-y)^{\alpha-1} & \text{for } y \in (0,1); \\ 0 & \text{otherwise.} \end{cases}$$

This is the p.d.f. of the Beta(β,α) distribution.

3.5 Let $\alpha > 0$.

(a) Show that $B(\alpha, 1) = \frac{1}{\alpha}$.

(b) Let $X \sim Be(\alpha, 1)$ distribution. Let $Y = \sqrt[r]{X}$ for some positive integer r. Show that Y also has a Beta distribution, and find its parameters.

Solution.

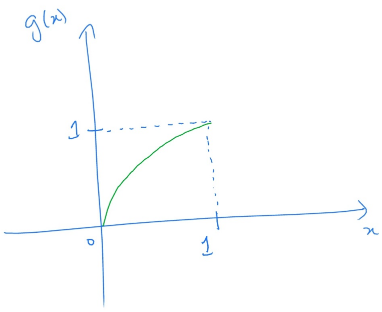
(a) From Lemma 2.2 we have $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$, hence

$$B(\alpha, 1) = \frac{\Gamma(\alpha)\Gamma(1)}{\Gamma(\alpha+1)} = \frac{1}{\alpha}.$$

(b) We aim to use Lemma 3.1. From (a) we have,

$$\begin{aligned} f_X(x) &= \begin{cases} \frac{1}{B(\alpha,1)} x^{\alpha-1} & \text{for } x \in (0,1), \\ 0 & \text{otherwise,} \end{cases} \\ &= \begin{cases} \alpha x^{\alpha-1} & \text{for } x \in (0,1), \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Hence, $R_X = (0, 1)$. The function $g(x) = \sqrt[r]{x}$



is strictly increasing on $(0, 1)$ and $g(R_X) = (0, 1)$. We have $g^{-1}(y) = y^r$ and $\frac{d}{dy} g^{-1}(y) = r y^{r-1}$. Hence, by Lemma 3.1, for $y \in (0, 1)$ we have

$$f_Y(y) = \alpha y^{r(\alpha-1)} \cdot r y^{r-1} = r\alpha y^{r\alpha-1}.$$

Thus,

$$f_Y(y) = \begin{cases} r\alpha y^{r\alpha-1} & \text{if } y \in (0,1) \\ 0 & \text{otherwise} \end{cases}$$

which is the p.d.f. of a $Be(r\alpha, 1)$ distribution. So Y has a Beta distribution with parameters $r\alpha$ and 1.

3.6 Let X have a uniform distribution on $[-1, 1]$. Find the p.d.f. of $|X|$ and identify the distribution of $|X|$.

Solution. If $x < 0$ then $\mathbb{P}[|X| \le 0] = 0$. And since $|X| \in [0, 1]$, if $x > 1$ we have $\mathbb{P}[|X| \le x] = 1$. For $x \in [0, 1]$ we have

$$\mathbb{P}[|X| \le x] = \mathbb{P}[-x \le X \le x] = \int_{-x}^{x} f_X(u)\, du = \int_{-x}^{x} \frac{1}{2}\, du = x.$$

Differentiating, we have

$$f_{|X|}(x) = \begin{cases} 1 & \text{for } x \in [0,1]; \\ 0 & \text{otherwise.} \end{cases}$$

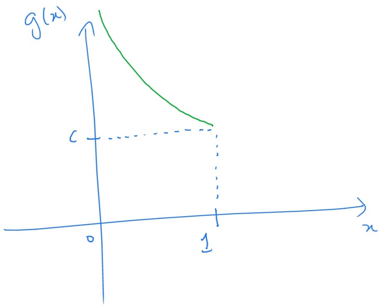
Therefore, $|X|$ is uniform on $[0,1]$.

Pitfall: If we try the 'standard' method of Lemma 3.1, we'll have $g(x) = |x|$, which is not monotone on $[-1,1]$; so the formula $f_Y(y) = f_X(g^{-1}(y))|\frac{dg^{-1}}{dy}|$ does not apply here – and it will give the wrong answer.

The same issue applies to several other questions in this section.

3.7 Let $\alpha, \beta > 0$ and let $X \sim Be(\alpha, \beta)$. Let $c > 0$ and set $Y = c/X$. Find the p.d.f. of Y.

Solution. We aim to use Lemma 3.1. We have $X \sim Be(\alpha, \beta)$ and $Y = c/X$, so set $g(x) = c/x$ where $g : (0,1) \to \mathbb{R}$.



Therefore, g is strictly decreasing on $(0,1)$. Then $g^{-1}(y) = \frac{c}{y}$, and $\frac{dg^{-1}}{dy} = \frac{-c}{y^2}$. Further, $f_X(x) > 0$ for $x \in (0,1)$, $R_X = (0,1)$ and $g(R_X) = (c, \infty)$. Hence, for $y > c$ we have

$$\begin{aligned} f_Y(y) &= \frac{1}{B(\alpha,\beta)} \left(\frac{c}{y}\right)^{\alpha-1} \left(1 - \frac{c}{y}\right)^{\beta-1} \frac{c}{y^2} \\ &= \frac{c^\alpha}{B(\alpha,\beta)} \frac{(y-c)^{\beta-1}}{y^{\alpha+\beta}}. \end{aligned}$$

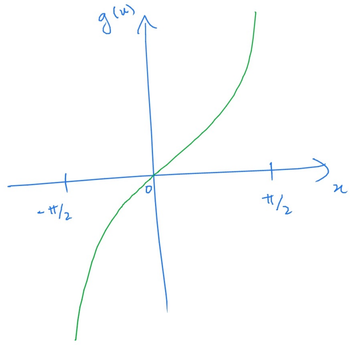
For $y \leq c$, we have $f_Y(y) = 0$.

3.8 Let Θ be an angle chosen according to a uniform distribution on $(-\frac{\pi}{2}, \frac{\pi}{2})$, and let $X = \tan\Theta$. Show that X has the Cauchy distribution.

Solution. We aim to use Lemma 3.1 (with $X = \Theta$). The random variable Θ has a $U(-\frac{\pi}{2}, \frac{\pi}{2})$ distribution, so

$$f_\Theta(\theta) = \begin{cases} \frac{1}{\pi} & \text{for } -\frac{\pi}{2} < \theta < \frac{\pi}{2}; \\ 0 & \text{otherwise.} \end{cases}.$$

The function $g(\theta) = \tan\theta$ is strictly increasing on the range of Θ, that is on $(-\frac{\pi}{2}, \frac{\pi}{2})$.



Moreover, $g^{-1}(y) = \arctan(y)$ and $g^{-1}(y) \in (-\frac{\pi}{2}, \frac{\pi}{2})$ for all $y \in \mathbb{R}$. Hence, for all $y \in \mathbb{R}$,

$$f_Y(y) = f_\Theta(g^{-1}(y)) \left| \frac{dg^{-1}}{dy} \right| = \frac{1}{\pi} \frac{1}{1+y^2}$$

Hence, Y has the Cauchy distribution.

3.9 Let X be a random variable with the p.d.f.

$$f(x) = \begin{cases} 1+x & \text{for } -1 < x < 0; \\ 1-x & \text{for } 0 < x < 1; \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability density functions of

(a) $Y = 5X + 3$

(b) $Z = |X|$

Solution.

(a) The function $g(x) = 5x + 3$ is strictly increasing, so we can use the formula

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

on each of the intervals in the definition of f_X. We note $g^{-1}(y) = \frac{y-3}{5}$ and $\frac{d}{dy} g^{-1}(y) = \frac{1}{5}$, so

$$\begin{aligned} f_Y(y) &= \begin{cases} \left(1 + \frac{y-3}{5}\right) \frac{1}{5} & \text{for } -1 < \frac{y-3}{5} < 0 \\ \left(1 - \frac{y-3}{5}\right) \frac{1}{5} & \text{for } 0 < \frac{y-3}{5} < 1 \\ 0 & \text{otherwise} \end{cases} . \\ &= \begin{cases} \frac{y+2}{25} & \text{for } -2 < y < 3 \\ \frac{8-y}{25} & \text{for } 3 < y < 8 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

(b) The function $g(x) = |x|$ is not monotonic, so instead we use that for $z \geq 0$

$$\begin{aligned}
F_Z(z) &= \mathbb{P}[Z \leq z] \\
&= \mathbb{P}[-z \leq X \leq z] \\
&= \begin{cases} \int_{-z}^{0}(1+u)\,du + \int_0^z (1-u)\,du & \text{for } z \leq 1 \\ 1 & \text{for } z > 1 \end{cases} \\
&= \begin{cases} 2z - z^2 & \text{for } z \leq 1 \\ 1 & \text{for } z > 1. \end{cases}
\end{aligned}$$

If $z < 0$ then $\mathbb{P}[|X| < z] = 0$. So the p.d.f. of Z is

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \begin{cases} 2(1-z) & 0 \leq z < 1 \\ 0 & \text{otherwise.} \end{cases}$$

3.10 Let X have the uniform distribution on $[a, b]$.

(a) For $[a, b] = [-1, 1]$, find the p.d.f. of $Y = X^2$.

(b) For $[a, b] = [-1, 2]$, find the p.d.f. of $Y = |X|$.

Solution.

(a) The probability density function of X is

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{for } x \in [-1, 1], \\ 0 & \text{otherwise.} \end{cases}$$

We have $Y \geq 0$, so $f_Y(y) = 0$ for $y \leq 0$. For $y \in (0, 1]$ we have

$$\mathbb{P}[Y \leq y] = \mathbb{P}[0 \leq X^2 \leq y] = \mathbb{P}[-\sqrt{y} \leq X \leq \sqrt{y}] = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}\,dy = \sqrt{y}.$$

Thus, $\mathbb{P}[Y \leq 1] = 1$ and hence $\mathbb{P}[Y \leq y] = 1$ for all $y \geq 1$. Differentiating, we obtain

$$f_Y(y) = \begin{cases} 0 & \text{for } y \leq 0, \\ \frac{1}{2} y^{-1/2} & \text{for } y \in (0, 1], \\ 0 & \text{for } y > 1. \end{cases}$$

(b) The probability density function of X is

$$f_X(x) = \begin{cases} \frac{1}{3} & \text{for } x \in [-1, 2], \\ 0 & \text{otherwise.} \end{cases}$$

We have $Y \geq 0$, so $f_Y(y) = 0$ for $y \leq 0$. For $y > 0$, we need to consider three cases; $y \in (0, 1]$ and $y \in (1, 2]$ and $y > 2$.
For $y \in [0, 1]$, we have

$$\mathbb{P}[Y \leq y] = \mathbb{P}[-y \leq X \leq y] = \int_{-y}^{y} \frac{1}{3}\,dy = \frac{2y}{3}.$$

For $y \in (1,2]$, we have

$$\mathbb{P}[Y \le y] = \mathbb{P}[Y \le 1] + \mathbb{P}[1 < X \le y] = \frac{2}{3} + \int_1^y \frac{1}{3}\,dy = \frac{2}{3} + \frac{y-1}{3}$$

For $y < 2$, we note that $\mathbb{P}[Y \le 2] = 1$ from the previous case, so $\mathbb{P}[Y \le y] = 1$ for all $y > 2$.

Differentiating, we obtain

$$f_Y(y) = \begin{cases} 0 & \text{for } y \le 0 \text{ or } y > 2, \\ \frac{2}{3} & \text{for } y \in [0,1], \\ \frac{1}{3} & \text{for } y \in (1,2]. \end{cases}$$

3.11 Let X have a uniform distribution on $[-1,1]$ and define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) = \begin{cases} 0 & \text{for } x \le 0; \\ x^2 & \text{for } x > 0. \end{cases}$$

Find the distribution function of $g(X)$.

Solution. Clearly $Y = g(X) \ge 0$, so $\mathbb{P}[Y \le y] = 0$ for all $y < 0$. We have

$$\mathbb{P}[Y = 0] = \mathbb{P}[X \le 0] = \int_{-1}^0 \frac{1}{2}\,dx = \frac{1}{2}.$$

For $y \in (0,1]$ we have

$$\mathbb{P}[Y \le y] = \mathbb{P}[Y = 0] + \mathbb{P}[0 < Y \le y] = \frac{1}{2} + \int_0^{\sqrt{y}} \frac{1}{2}\,dy = \frac{\sqrt{y}+1}{2},$$

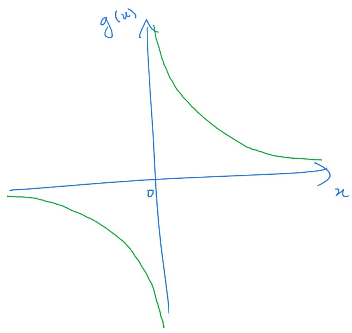
which means that $\mathbb{P}[Y \le 1] = 1$ and hence $\mathbb{P}[Y \le y] = 1$ for all $y \ge 1$.

To sum up,

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0, \\ \frac{1}{2} & \text{for } y = 0, \\ \frac{\sqrt{y}+1}{2} & \text{for } y \in (0,1], \\ 1 & \text{for } y > 1. \end{cases}$$

3.12 Let X be a random variable with the Cauchy distribution. Show that X^{-1} also has the Cauchy distribution.

Solution. The random variable X has p.d.f. $f_X(x) = \frac{1}{\pi(1+x^2)}$. Define $g : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$ by $g(x) = 1/x$.



Note that $\mathbb{P}[X = 0] = 0$, so the distribution of $Y = 1/X$ is still well defined. We have $Y = g(X)$, but g is not strictly monotone (for example, $g(-1) < g(1) > g(2)$).

We plan to show that $\mathbb{P}[Y \le y] = \mathbb{P}[X \le y]$ for all $y \in \mathbb{R}$, which means that X and Y have the same distribution function; hence the same distribution.

Let us look at the case $y = 0$ first. There is no value of $x \in \mathbb{R}$ such that $g(x) = 0$, so $\mathbb{P}[Y = 0] = 0$. Since X is a continuous distribution, $\mathbb{P}[X = 0] = 0$. Also $X < 0$ if and only if $Y < 0$, so we have $\mathbb{P}[X \le 0] = \mathbb{P}[X < 0] = \mathbb{P}[Y < 0] = \mathbb{P}[Y \le 0]$.

For $y < 0$, we have

$$\begin{aligned}\mathbb{P}[Y \le y] &= \mathbb{P}\left[\tfrac{1}{y} \le X < 0\right] \\ &= \int_{\frac{1}{y}}^{0} \frac{1}{\pi(1+x^2)}\,dx \\ &= \int_{y}^{-\infty} \frac{1}{\pi(1+(1/v)^2)}\,\frac{-1}{v^2}\,dv \\ &= \int_{-\infty}^{y} \frac{1}{\pi(1+v^2)}\,dv = \mathbb{P}[X \le y].\end{aligned}$$

Here, we substitute $x = 1/v$. For $y > 0$, we must split into two cases, giving

$$\begin{aligned}\mathbb{P}[\tfrac{1}{X} \le y] &= \mathbb{P}[X < 0] + \mathbb{P}[\tfrac{1}{y} \le X] \\ &= \mathbb{P}[X < 0] + \int_{\frac{1}{y}}^{\infty} \frac{1}{\pi(1+x^2)}\,dx \\ &= \mathbb{P}[X < 0] + \int_{y}^{0} \frac{1}{\pi(1+(1/v)^2)}\,\frac{-1}{v^2}\,dv \\ &= \mathbb{P}[X < 0] + \int_{0}^{y} \frac{1}{\pi(1+v^2)}\,dv \\ &= \mathbb{P}[X < 0] + \mathbb{P}[0 \le X \le y] = \mathbb{P}[X \le y].\end{aligned}$$

Again, we substitute $x = 1/v$.

Hence, $\mathbb{P}[X \le y] = \mathbb{P}[Y \le y]$ for all $y \in \mathbb{R}$. Since X has a Cauchy distribution, so does Y.

Challenge Questions

3.13 If we were to pretend that $g(x) = 1/x$ was strictly monotone, we could (incorrectly) apply Lemma 3.1 and use the formula $f_Y(y) = f_X(g^{-1}(y))|\frac{dg^{-1}}{dy}|$ to solve **3.12**. We would still arrive at the correct answer. Can you explain why?

Can you construct another example of a case in which the relationship $f_Y(y) = f_X(g^{-1}(y))|\frac{dg^{-1}}{dy}|$ holds, but where the function g is not monotone?

Hint. Carefully examine the proof of the formula for $f_Y(y) = f_X(g^{-1}(y))|\frac{dg^{-1}}{dy}|$ when g *is* strictly monotone, and compare it to the solution of **3.12**.

In general though, if g is not strictly monotone, the formula will not work!

3.14 Let Y and α, β be as in Question **3.7**.

(a) If $\alpha > 1$, show that $\mathbb{E}[Y] = \frac{c(\alpha+\beta-1)}{\alpha-1}$.

(b) If $\alpha \le 1$ show that $\mathbb{E}[Y]$ is not defined.

Solution.

(a) We use that $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$. So

$$\begin{aligned}
\mathbb{E}\left[\frac{c}{X}\right] &= \int_{-\infty}^{\infty} \frac{c}{x} f_X(x)\, dx \\
&= \int_0^1 \frac{c}{x} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}\, dx \\
&= \frac{c}{B(\alpha,\beta)} \int_0^1 x^{\alpha-2}(1-x)^{\beta-1}\, dx \\
&= \frac{cB(\alpha-1,\beta)}{B(\alpha,\beta)},
\end{aligned}$$

Expanding the Beta functions here in terms of the Gamma function, and using that $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$, we get

$$c\frac{\frac{\Gamma(\alpha-1)\Gamma(\beta)}{\Gamma(\alpha+\beta-1)}}{\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}} = c\frac{\frac{\Gamma(\alpha-1)\Gamma(\beta)}{\Gamma(\alpha+\beta-1)}}{\frac{(\alpha-1)\Gamma(\alpha-1)\Gamma(\beta)}{(\alpha+\beta-1)\Gamma(\alpha+\beta-1)}} = \frac{c(\alpha+\beta-1)}{\alpha-1}.$$

(b) If we try to calculate $\mathbb{E}[Y]$ then, using the p.d.f. obtained in **3.7**(b) we want

$$\frac{c^\alpha}{B(\alpha,\beta)} \int_c^\infty \left(\frac{y-c}{y}\right)^{\beta-1} y^{-\alpha}\, dy. \tag{3.1}$$

We need to show that the integral does not converge. The idea is to note that $y - c \approx y$, for large y, which means that $(\frac{y-c}{y})^{\beta-1} \approx 1$; leaving us with $\int_c^\infty y^{-\alpha}\, dy$ which diverges to ∞.

To implement this idea, we could begin by noting that

$$(3.1) \ge \frac{c^\alpha}{B(\alpha,\beta)} \int_{c+1}^\infty \left(\frac{y-c}{y}\right)^{\beta-1} y^{-\alpha}\, dy \tag{3.2}$$

Since $y > c+1$, we have $\frac{1}{c+1} \le \frac{y-c}{y} \le 1$. If $\beta \ge 1$ then $(\frac{y-c}{y})^{\beta-1} \ge \frac{1}{(c+1)^{\beta-1}}$, and if $\beta \in (0,1]$ then $(\frac{y-c}{y})^{\beta-1} \ge 1$. Hence,

$$(3.2) \ge \frac{c^\alpha}{B(\alpha,\beta)} \min\left(\frac{1}{(c+1)^{\beta-1}}, 1\right) \int_{c+1}^\infty y^{-\alpha}\, dy.$$

Since $\int_{c+1}^\infty y^{-\alpha}\, dy$ diverges for $\alpha \le 1$, we have that $\mathbb{E}[Y]$ does not exist for such α.

4 Multivariate Distribution Theory

Warm-up questions

4.1 Let $T = \{(x, y) \ : \ 0 < x < y\}$. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x,y) = \begin{cases} e^{-2x-y} & \text{for } (x,y) \in T; \\ 0 & \text{otherwise.} \end{cases}$$

Sketch the region T. Calculate $\int_0^\infty \int_0^\infty f(x,y)\,dx\,dy$ and $\int_0^\infty \int_0^\infty f(x,y)\,dy\,dx$, and verify that they are equal.

Solution. See **4.2** for a sketch of T. We have

$$\begin{aligned}
\int_{y=0}^\infty \int_{x=0}^\infty f(x,y)\,dx\,dy &= \int_{y=0}^\infty \int_{x=0}^y e^{-y}e^{-2x}\,dx\,dy = \int_0^\infty e^{-y}\left[\frac{-1}{2}e^{-2x}\right]_{x=0}^y dy \\
&= \int_0^\infty e^{-y}\left(\frac{-1}{2}e^{-2y} + \frac{1}{2}\right) dy = \frac{1}{2}\int_0^\infty e^{-y} - e^{-3y}\,dy \\
&= \frac{1}{2}\left[-e^{-y} + \frac{1}{3}e^{-3y}\right]_{y=0}^\infty = \frac{1}{2}\left(1 - \frac{1}{3}\right) = \frac{1}{3}
\end{aligned}$$

and

$$\begin{aligned}
\int_{x=0}^\infty \int_{y=0}^\infty f(x,y)\,dy\,dx &= \int_{x=0}^\infty \int_{y=x}^\infty e^{-y}e^{-2x}\,dx\,dy = \int_0^\infty e^{-2x}\left[-e^{-y}\right]_{y=x}^\infty dx \\
&= \int_0^\infty e^{-2x}e^{-x}\,dx = \int_0^\infty e^{-3x}\,dx = \left[\frac{-1}{3}e^{-3x}\right]_{x=0}^\infty = \frac{1}{3},
\end{aligned}$$

which are equal.

Pitfall: We must be careful to get the limits on the integrals correct. For $\int\int \dots dy\,dx$, we first allow x to vary between $0 \dots \infty$, which means that to cover T we must allow y to vary between x and ∞. We think of covering T by vertical lines, one for each $x = 0 \dots \infty$, each line with constant x and with y ranging from x up to ∞. It's helpful to draw a picture.

Alternatively, for $\int\int \dots dx\,dy$, we first allow y to vary between 0 and ∞, which means that to cover T we must allow x to vary between 0 and y. We think of covering T by horizontal lines, one for each $y = 0 \dots \infty$, each line with constant y and with x ranging from 0 up to y.

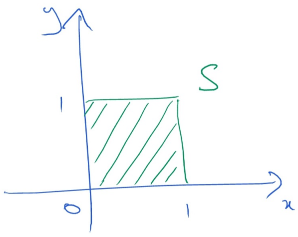
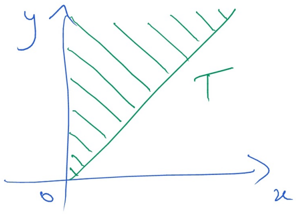
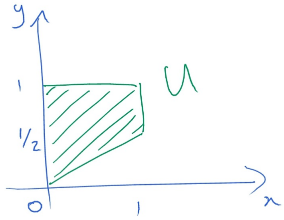
4.2 Sketch the following regions of $\mathbb{R}^2$.

(a) $S = \{(x, y) \ : \ x \in [0, 1],\ y \in [0, 1]\}$.

(b) $T = \{(x, y) \ : \ 0 < x < y\}$.

(c) $U = \{(x, y) \ : \ x \in [0, 1],\ y \in [0, 1],\ 2y > x\}$.

Solution.


y
S
1
0
1
x

y
T
0
x

y
U
1
1/2
0
1
x